# The Atoms of Person
## Limitations on Concept Formation

Jolijn Sonnaert

Even though person markers such as agreement suffixes, clitics or independent pronouns are very common crosslinguistically, many questions about the possible person referents these markers can refer to remain. Consider for example the inclusive, which refers to the referents of both first and second person, respectively speaker (called *i* in this paper) and hearer (called *u*): *iu*. In a representative sample of 370 languages, the other logically possible person referent combinations, *io* (for speaker and other) and *uo* (for hearer and other), are unattested and therefore impossible person distinctions in natural language. The literature often equates *io* and *uo* with first and second person plural, respectively. However, I show that pronominal plural is not an instance of '+ third person' or '+ *o*', but that it is in fact an instance of '+ associates' or '+ *a*', resulting in the plurals *iua*, *ia*, *ua* and *oa*. The CONCEPT FORMATION CONSTRAINT in the kite framework (see Jaspers 2012; Seuren & Jaspers 2014; Roelandt 2016) makes generalisations about lexicalisation in closed lexical fields and predicts exactly the gap in lexicalisation we see for person, namely the absence of person markers for the combinations *io* and *uo*. As such, the fact that these are impossible person distinctions is part of a broader generalisation on lexicalisation in natural language.

### *1. Introduction*

Pronominal paradigms are typically analysed as having three person atoms, corresponding to speaker, hearer and other and being referred to by first, second and third person.[1] In this paper, I refer to these atoms as *i*, *u* and *o* respectively, following a.o. Harbour (2016) and Ackema & Neeleman (To appear). As for plural, I argue that pronominal plural is formed by adding

[1]Note that these atoms and their combinations are not the same as the morphosyntactic features like [± participant], [± speaker], etc. familiar from the literature (e.g. Harley & Ritter 2002; Bobaljik 2008; Harbour 2016). Instead, these atoms and their combinations are the referents that these features can select for pronominals to refer to. This paper is only concerned with the atoms and combinations, not with the morphosyntactic features.

*Proceedings of ConSOLE XXV, 2017, 44–68*
http://www.sole.leidenuniv.nl

associates (*a*) to these person atoms (as already suggested by Bobaljik (2008) and Ackema & Neeleman (To appear)) rather than by adding third person, as commonly assumed in a lot of the literature on person (e.g. Cysouw 2003; Harbour 2016). The distinction between associates and others is a crucial one, and will be motivated in Section 3.1.

The relationship between possible person atoms and morphological person markers is not one to one, since on top of first, second and third person, languages can also have a distinct inclusive person. The inclusive refers to a group consisting of both the speaker and the hearer and is commonly analysed as a combination of those. For example, Cysouw (2003) describes the inclusive as '1+2', Harbour (2016) and Ackema & Neeleman (To appear) as '*iu*'. However, languages have different ways of expressing this inclusive. Consider for example Tümpisa Shoshone below, where inclusive *tammü* refers to a *we* including the hearer, as opposed to *nümmü* referring to a *we* excluding the hearer. This latter one is also called the *exclusive* or the first person plural.

(1) Tümpisa Shoshone (Dayley 1979)

| | SG | | | PL | | |
|---|---|---|---|---|---|---|
| incl | | | | *ta-mmü* | iu(a) | (we, including you) |
| 1 | *nü* | i | (I) | *nü-mmü* | ia | (we, but not you) |
| 2 | *ü* | u | (you) | *mü-mmü* | ua | (you) |
| 3 | (demonstr) | o | (he, she, it) | (demonstr) | oa | (they) |

Languages like Tümpisa Shoshone, which have one distinct pronominal to express this combination of speaker and hearer, are called *inclusive languages* (terminology from Daniel 2005), as opposed to *non-inclusive languages* such as English where the inclusive is expressed syncretically with the exclusive.

The cell for pronouns that would be both 'inclusive' and 'singular' in the paradigm in (1) is empty because the inclusive can never be singular. It necessarily refers to both the speaker and hearer. Also, in agglutinative paradigms where person and number are expressed by separate morphemes like Tümpisa Shoshone above, the inclusive takes plural rather than singular morphology: just like the other plural pronouns, it adds *-mmü*.

This is not to say the inclusive cannot have a number distinction. Some languages, i.e. *minimal-augmented languages* like Rembarrnga below, have two pronominals to refer to a combination of speaker and hearer. They distinguish a minimal from an augmented inclusive: that is, an inclusive that consists of the minimal amount of participants needed for the inclusive person (*i* and *u*), *yu̶kku̶*, and an inclusive that adds to this associates (*i*, *u* and *a*), *ngakorrbbarrah*. Note that just like the plural as opposed to the singular in Tümpisa Shoshone, the augmented persons have a number morpheme, *-barrah*, that the minimal do not.

(2) Rembarrnga (Cysouw 2003)

| | MIN | | | AUGM | | |
|---|---|---|---|---|---|---|
| incl | *yu̶kku̶* | iu | (we: me & you) | *ngakorr-bbarrah* | iua | (we: me, you, & more) |
| 1 | *ngu̶nu̶* | i | (I) | *yarr-bbarrah* | ia | (we, but not you) |
| 2 | *ku̶* | u | (you) | *nakorr-barrah* | ua | (you) |
| 3 | *nawu̶/ ngadu̶* | o | (he, she, it) | *barr-barrah* | oa | (they) |

To summarise: we can distinguish three types of languages when considering the inclusive person: non-inclusive languages like English that express the inclusive syncretically with the exclusive, inclusive languages like Tümpisa Shoshone with an inclusive that behaves like the plural persons, and minimal-augmented languages with an inclusive that behaves like the singular other persons, i.e. minimal, and one that behaves like the plural other persons, i.e. augmented. Considering that in addition to the lexicalisations of the single person atoms (i.e. first, second and third person markers), languages can also distinguish a combination of two of these atoms (inclusive, *iu*), the question arises whether they can lexicalise even more person distinctions. The three person atoms can be combined in eight different ways (3).[2]

(3)
| | | | |
|---|---|---|---|
| a. | (ø, expletives) | e. | *iu*, inclusive |
| b. | *i*, first person | f. | *io*, – |
| c. | *o*, second person | g. | *uo*, – |
| d. | *u*, third person | h. | (*iuo*, impersonals) |

I use a sample of 370 languages to show that only six out of these eight possibilities can be lexicalised as simplex person morphemes: *io* and *uo* cannot be lexicalised. This is especially clear when also considering number. For a mininal-augmented language as discussed above, the six logically possible person distinctions in the two numbers would yield a 12-cell paradigm, however, for these languages only 8-cell paradigms like the one for Rembarrnga are attested, with both numbers distinct for *i, u, o* and *iu*.

The absence of exactly *io* and *uo* in person lexicalistion is predicted in the kite framework developed by Seuren & Jaspers (2014). Seuren & Jaspers look at concept formation in closed lexical fields and the logical relations that hold between the concepts. In doing so, the framework makes predictions about lexicalisation. For example, for quantifiers, it predicts that languages do not have a simplex morpheme to express the contradiction of *all*, namely *not all*, nor to express the disjunction *all or no*. This is the CONCEPT FORMATION CONSTRAINT (CFC). I will show that the same constraint applies to person.

The paper is organised as follows: I first discuss the person lexicalisations and the unlexicalisable combinations in detail in Section 3. Then I introduce the kite framework and explain how it makes predictions about lexicalisation in Section 4. In Section 5, I show how this framework predicts exactly and only the attested person distinctions and the relations that hold between them as part of a broader generalisation on lexicalisation.

## 2. *The sample*

The sample collected consists of two parts. The first part is a small sample of independent pronouns for only 30 languages, collected to be representative for the genetic and geographic variation as discussed in *The World Atlas of Language Structures* (Dryer & Haspelmath 2013). This was then extended to a sample of 370 languages with all paradigms available in the typologies and analyses of a.o. Forchheimer (1953); Harley & Ritter (2002); Baerman et al. (2005);

[2]This paper does not discuss the absence and presence of all atoms, i.e. *ø* and *iuo*. These might correspond to expletive pronouns (ø, not referring to any person atoms) and impersonal pronouns (*iuo*, referring to all person atoms) as proposed by Ackema & Neeleman (To appear). This paper focuses only on the referential and personal use of person morphemes: the options in (3b)-(3g).

Daniel (2005); Bobaljik (2008); Cysouw (2003); Harbour (2016); Ackema & Neeleman (To appear). The sample now includes all kinds of person markers: free pronouns, clitics and agreement markers.

## *3. The lexicalisation of person*

Even though analyses of person differ on which morphosyntactic features arrange the paradigms of pronominals, all these accounts assume that person markers can refer to three different persons: a speaker *i*, a hearer *u* and an other *o*. Pronominals can refer to these single atoms with first, second and third person, or to a combination of atoms *iu* with the inclusive. In this section, I will discuss the lexicalisations of all person atoms and their combinations and show that of the latter, only the inclusive ever receives a simplex lexicalisation in language, meaning that only the inclusive can be expressed with a single person morpheme. The other two combinations (*io* and *uo*) never have a distinct person morpheme.

I start with discussing the distinction between *o*, third person, and *a*, plural, in Section 3.1 on pronominal number. Next, I discuss the person atom lexicalisations in detail in Section 3.2, to conclude this section with a note on syncretism in Section 3.3.

### *3.1. Number*

Looking at number, many analyses of person collapse third person and plural (e.g. Cysouw 2003; Moskal 2014; Harbour 2016). Consider for example the following terminology used in Cysouw (2003) to represent the person and number paradigm, where the plural persons are marked as the relevant person '+ third person'.[3]

(4) The person and number paradigm in Cysouw (2003)

| PERSON | SG | PL |
|---|---|---|
| inclusive | | 1+2(+3) |
| exclusive | 1 | 1+3 |
| addressee | 2 | 2+3 |
| other | 3 | 3+3 |

The same goes for Harbour's (2016) person analysis, for which the ontology is shown in Fig. 1. In his system, *i* represents the (unique) speaker, *u* represents the (unique) hearer and the *o*s represent an infinite amount of others. According to Harbour, applying the features he proposes to this ontology derives all persons attested, described in (5). The subscript *o* points to the possible addition of one or more 'others' for plural, the normal case *o* points to one 'other' for third person.

[3] I do not wish to imply that Cysouw is committed to the belief that third person is plural, but the notation used is fairly widespread and might invite this conclusion.

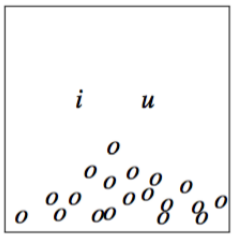


*Figure 1*: *The person and number ontology in Harbour (2016).*

(5) a. $iu_o$ = inclusive
b. $i_o$ = first person
c. $u_o$ = second person
d. $o_o$ = third person

The fact that both plural and third person consist of this *o* is a defining characteristic of Harbour's system.

If plural and third person were the same, *uo* and *io* would refer to first and second person plural. However, the morphology and semantics of third person and plural show that it is incorrect to conflate both. In what follows, I first discuss their morphology (Section 3.1.1). I show that no language uses the same morpheme for both plural and third person, which would not be unexpected if they were indeed the same. For semantics, I follow Ackema & Neeleman (To appear) in that the reference of a plural pronoun cannot include a third person. I demonstrate this with the results of a small questionnaire in Dutch, and show that the generalisation holds for all persons (Section 3.1.2). Instead, plural should be analysed as '+ associates' (*a*) (as already suggested by a.o. Bobaljik (2008) and Ackema & Neeleman (To appear)), which is distinct from third person (*o*).

### *3.1.1. Morphology of plural and third person*

If third person and plural were indeed the same, we would expect languages to exist with morphologically transparent paradigms (e.g. agglutinative languages) where plural and third person are expressed by the same morpheme, as in the following paradigm:[4]

(6) Hypothetical paradigm for 'plural = +3$^{rd}$ person':

| PERS | SG | PL |
|---|---|---|
| iu | | $\alpha$(-$\delta$) |
| i | $\beta$ | $\beta$-$\delta$ |
| u | $\gamma$ | $\gamma$-$\delta$ |
| o | $\delta$ | $\delta$(-$\delta$) |

[4]The claim that third person and plural are the same is more often than not implicit in the literature, and is reflected in a type of notation that represents plural pronouns as '1+3', '2+3', etc. (e.g. Cysouw 2003), or *io*, *uo*, etc. (e.g. Harbour 2016). This does not necessarily mean that the relevant authors are committed to the belief that the third person and plural are identical. For this reason, I focus on Forchheimer (1953), since he has made the claim in the most explicit form that I am aware of. His claim is moreover formulated at a very general level, and ignores the distinction between semantic atoms (referents) and morphological atoms (features). My argument against the claim is based on a version that would assume that not only featural but also semantic identity will be reflected in a formal identity of some sort, and explores the consequences of such an assumption.

However, in the 370 languages from the sample, this pattern does not occur.

The only defense of the claim that 'plural = +3$^{rd}$ person' in terms of the morphology of attested paradigms is Forchheimer (1953), in which some languages are described as having the third person marker as a plural marker. However, for each of these languages, a more correct reanalysis shows that these are not true examples of 'plural = +3$^{rd}$ person' languages. I demonstrate this by discussing one of these composite forms, the independent pronouns of Kalaw Lagaw Ya (Pama-Nyungan). Similar reanalyses are available for the other languages, Arrernte (Pama-Nyungan), Siuslaw, Coos and Chinook (Penutian) and Hurrian and Sumerian; however, discussing all of them is beyond the scope of this paper.

Forchheimer mentions the following links between third person and plural in Kalaw Lagaw Ya (KLY) (Forchheimer 1953:127-128):

(7) a. The exclusive plural *ngoi* 'may be a combination of [first person singular] *ngai*. . . and [third person singular] *noi*' (Forchheimer 1953:127).
b. *Tana*, third person plural, is used as a plural suffix in the second person *ngitana* (Forchheimer 1953:127-128).

The pronoun paradigm given by Forchheimer is represented in (8) with the pronouns under discussion boldfaced.

(8) KLY personal pronouns by Forchheimer (1953:127)

| | SG | PL |
|---|---|---|
| iu | | *ngalpa* |
| i | ***ngai*** */ nazo* | ***ngoi****, ngöi* |
| u | *ngi* | *ngi****ta(na)*** |
| o | ***noi****, nu / na, nadu* | ***tana****, ari* |

If exclusive were a clear example of a '+ 3$^{rd}$ person plural' as suggested in (7a), we would expect to see *ngainoi* as the plural of first person, with the third person singular *noi* added to the first person singular *ngai*. This is not the case, so the link suggested by Forchheimer is more subtle: he seems to suggest that a part of the third person singular, *-oi*, is added to a part of the first person singular, *ng-*. Below, I show how other, more recent, analyses of KLY show no such link between the exclusive and third person singular, and that second person plural is formed by adding a distinct plural marker to second person singular, rather than adding third person plural to second person singular as suggested in (7b).

In Ray's description of KLY in (9) (1907:22), both the exclusive and the second person plural have slightly different forms. The same goes for the descriptions of Ford & Ober (2004:138) and Round & Stirling (2015:264), shown in (10).

(9) KLY pronouns by Ray (1907:22)

| | SG | PL |
|---|---|---|
| iu | | *ngalpa* |
| i | *ngai* | ***ngoi*** |
| u | *ngi* | ***ngita*** |
| o | ***nui*** */ na* | ***tana*** |

(10) KLY pronouns by Ford & Ober (2004:138) and Round & Stirling (2015:264)

| | SG | PL |
|---|---|---|
| iu | | *ngalpa* |
| i | *ngay* | ***ngoey*** |
| u | *ngi* | ***ngitha*** |
| o | ***nuy*** */ na* | ***thana*** |

The link Forchheimer mentions between exclusive and third person singular (7a) does not hold. As opposed to Forchheimer's *noi* for third singular, Ray notes *nui* and Ford & Ober and Round & Stirling note *nuy* as third person singular. Their exclusive *ngoi* and *ngoey* do not show a clear relation to third singular.[5]

As for (7b), the second person plural pronoun in (9) and (10) lacks the optional *-na* that Forchheimer analyses as a part of the third person plural pronoun *tana*. As such, second person plural *ngita* is a combination of second person singular *ngi* and *-ta*, the latter which Ray analyses as a plural marker. Support for Ray's claim comes from the KLY deictic system discussed by Ford & Ober and shown in (11), where *-tha* (or in Ray's paradigm, *-ta*) does indeed occur as a plural marker. As such, *ngi-t(h)a* should be decomposed as 2sg+pl and *t(h)a-na* as pl+3sg.[6]

(11) The KLY deictic system (with more distinctions in Ford & Ober 2004:137)

| | SG MASC | SG FEM | DU | PL |
|---|---|---|---|---|
| proximate | *in* | *ina* | *ipal* | *i**tha*** |
| remote | *senaw* | *sena* | *sepal* | *se**tha*** |
| 'over there' | *pinungap* | *pinangap* | *pipalngap* | *pi**tha**ngap* |
| 'up there' | *pinuka* | *pinaka* | *pipalka* | *pi**tha**ka* |
| 'down there' | *pinuguy* | *pinaguy* | *pipalguy* | *pi**tha**guy* |
| 'up at the front' | *pinupay* | *pinapay* | *pipalpay* | *pi**tha**pay* |
| 'down at the back' | *pinupun* | *pinapun* | *pipalpun* | *pi**tha**pun* |

In Ray (1907); Ford & Ober (2004); Round & Stirling (2015), there is no morphological link between exclusive and third person, nor between second person plural and third person. Therefore, KLY is not a convincing example of a language that uses the third person morpheme to express plural on the other persons.

Forchheimer's alleged examples of 'plural = +$3^{rd}$ person' languages are far from convincing (1953). Taken together with the absence of the paradigm in (6) in the 370 languages I looked at, we can thus conclude that it is not a possible paradigm.

The next section explains how semantics and more specifically the reference of plural pronouns confirms this.

### *3.1.2. Semantics of plural and third person*

In this section, I show that plural pronouns cannot refer to a group that adds 'others' or third persons to the relevant person. As such, *io* and *uo* do not correspond to respectively first and second person plural, but are unlexicalisable combinations, as further argued for in Section 3.2.

Ackema & Neeleman (To appear) also argue for a distinction between plural and third person. They distinguish two atoms in the person ontology: associates (*a*), being the non-focal referents of a plural pronominal, and others (*o*), being the third person referent.[7] They impose the same restriction as discussed in this paper: the atoms *i* and *u* cannot occur with *o*. Instead, these are pluralised by adding *a*. However, they have to stipulate this, as they do not explain why

[5]How the exclusive is to be decomposed remains unclear.

[6]It is unclear why plural is a suffix on second person and a prefix on third person. However, it is not uncommon for third person to behave differently from the other persons.

[7]In fact, they introduce two different kinds of associates: $a_i$ and $a_u$, associates to the speaker and associates to the hearer. However, nothing in their system relies on this distinction and $a_i$ and $a_u$ can easily be conflated into one kind of associate.

*o* cannot occur with *i* or *u*.[8] The analysis introduced in Section 5, on the other hand, shows this cooccurrence restriction to be part of a generalisation on concept formation that holds for lexicalisation in general. Also, they do not extend this restriction to third person, which they analyse as one *o* for third person singular and several *o*s for third person plural. However, I argue that plural is formed by adding associates for all persons.

The argument they use to distinguish *a* from *o* is of a semantic nature:

> [W]hat is an *o* at a particular point in the discourse cannot be included in the reference of a first or second person plural pronoun without first being turned into an associate in some way. The other side of the coin is that an associate cannot be treated as an *o* without first being introduced as such (Ackema & Neeleman To appear:71).

This is illustrated with the example in (12): answer a. is considered odd because at this point in the conversation, George Clooney and Ad have not been established as associates. In this context, answer b. would be better. If a. is indeed uttered, then this has the comical effect of implying that George Clooney and Ad are in fact close friends.

(12) (Peter:) Do you know whether George Clooney likes good coffee?
  a. (Ad:) #Yes, we both drink Illy.
  b. (Ad:) Yes, he drinks Illy, just like me.

However, after uttering b., an associate relation is established, namely the one of drinking Illy. This means that from this point on in the conversation, Ad can use a plural pronoun to refer to himself and George Clooney. For example:

(13) (Following to the context in (12b)) We know good coffee when we see it.

This suggests that a plural pronoun is not made up of '+ third person' but rather of '+ associates' and that third person and associates are distinct elements.

In order to check whether this effect corresponds to speakers' judgements, I have conducted a small survey including 32 speakers of Flemish Dutch. The informants had to rate the suitability of the answers in situations such as the one described in Ackema & Neeleman (To appear), shown above, on a scale of 0 to 5 with 0 being a very odd and unexpected answer and 5 being a perfectly suitable answer. The questions were randomised. I briefly discuss the results for the example in (12) in (14)-(16), and afterwards, show that the same effect holds for third person in (17)-(19).

For each person discussed, I provide examples with and without an associate relation, and with a plural pronoun or a combination of two singular ones. For ease of exposition, I have underlined the third person under discussion and boldfaced the pronoun that is meant to refer to this. The example sentences discussed here are given in English.

[8]The non-occurrence of *o* with either *i* or *u* follows from the structure of their input set, which is a set of nested structures on which the person features operate. *o* cannot occur with *i* or *u* because their nested structures cannot be selected by these features to derive plural pronouns. However, they provide no motivation as to why this input set is structured in this particular way.

(14) **First person, no associate relation:**
Peter and Ad are watching a movie starring George Clooney. Peter asks Ad: 'Do you know whether George Clooney likes to drink coffee?' to which Ad replies:
a. 'Yes, **we** both like to drink Nespresso.'
b. 'Yes, **he** likes to drink Nespresso, just like me.'

(15) **First person, associate relation:**
Peter and Ad are watching a movie starring George Clooney. Ad is telling Peter that he and George Clooney go for a walk together every Sunday and that they have been friends for years now. Peter asks Ad: 'Do you know whether George Clooney likes to drink coffee?' to which Ad replies:
a. 'Yes, **we** both like to drink Nespresso.'
b. 'Yes, **he** likes to drink Nespresso, just like me.'

The table in (16) shows the average score and standard deviation for each example, with the score boldfaced for the distinctions relevant to the issue at hand.[9]

(16) Judgements for first person plural

| Example | Associate relation | Number of the pronoun | Response (0-5) | SD |
|---|---|---|---|---|
| (14a) | x | pl | **2.44** | 1.58 |
| (14b) | x | sg | 4.18 | 0.95 |
| (15a) | ✓ | pl | **3.38** | 1.44 |
| (15b) | ✓ | sg | 4.41 | 0.74 |

The response in the sentence without an associate relation was rated considerately better for the singular pronouns (14b) than for the plural pronoun (14a): 4.22 versus 2.45. This confirms that a third person (i.e. George Clooney) cannot be included in the reference of a first person plural pronoun. Interestingly, when establishing an associate relation as in (15), the response with the plural pronoun is now rated significantly better (with a p-value of 0.001): 3.39 in (15a) for *we* compared to 2.45 in (14a) without the associate relation. However, even in the sentence with the associate relation, the first person plural pronoun is still rated less good than the singular pronouns (with a p-value of $< 0.001$): 3.39 versus 4.42. This is because an associate relation is not entirely fixed and can be perceived differently by different people to a certain degree.

According to Ackema & Neeleman, only a first and second person plural pronoun cannot include reference to third person. However, a similar example for third person shows the same effect.[10] See for example (17) without and (18) with a possible associate relation. In both cases the plural pronoun is rated less good when it includes the third person (in this situation: Julia Roberts) but it is considered better when that person is an associate, as the numbers in (19) show.

[9]I would like to thank Cora Pots for her help with the statistics of this survey.

[10]I am discussing the animate use of third person pronouns. Third person pronouns referring to inanimates such as 'Have you seen the spoons? **They** are not in the drawer' have a homogenous plural just like the nouns they refer to, and are not discussed in this paper.

(17) **Third person, no associate relation:**
Peter and Ad are watching a movie starring George Clooney and Julia Roberts. Peter asks Ad: 'Do you know whether George Clooney likes to drink coffee?' to which Ad replies:

a. 'Yes, **they** both like to drink Nespresso.'
b. 'Yes, he likes to drink Nespresso, just like **Julia Roberts**.'

(18) **Third person, associate relation:**
Peter and Ad are watching a movie starring George Clooney and Julia Roberts. Peter asks Ad: 'Don't you think Julia Roberts and George Clooney act so well together? . . . By the way, do you know whether George Clooney likes to drink coffee?' to which Ad replies:

a. 'Yes, **they** both like to drink Nespresso.'
b. 'Yes, he likes to drink Nespresso, just like **Julia Roberts.**"

The table below shows the scores for the different answers.

(19) Judgements for third person plural

| Example | Associate relation | Number of the pronoun | Response (0-5) | SD |
|---|---|---|---|---|
| (17a) | x | pl | **1.79** | 1.29 |
| (17b) | x | sg | 4.15 | 1.18 |
| (18a) | ✓ | pl | **2.31** | 1.40 |
| (18b) | ✓ | sg | 3.62 | 1.33 |

Just like the numbers for first person have shown above, the score for the answer with the plural is significantly better with the associate relation (18a) than without (17a) (with a p-value of 0.020).[11] This confirms that third person plural includes reference to associates rather than to third persons just like first and second person do.

Even though the associate relation discussed by Ackema & Neeleman is subject to some personal variation (in the sense that what counts as an associate relation for some may not for others) the results of the survey confirm that an other (*o*, third person) cannot be included in the reference of a plural pronoun without first being turned into an associate. This goes for all person distinctions in Dutch. As such, plural is formed by adding associates and *io* and *uo* do not refer to a plural first or second person. This will be relevant in what follows, where I discuss the lexicalisations or lack thereof of all logically possible person atoms and combinations.

### *3.2. Person atoms and their combinations*

Consider again the person atoms and their logically possible combinations:[12]

[11] The answer with the singular pronoun in the situation with the associate relation, (18b), is rated less good than the corresponding answer without the associate relation in (17b). At this point it is unclear why this is, although it may be related to the repetition of the proper name in (18a).

[12] I assume that the referent *a* is not a person atom, since it differs from *i, u* and *o* in crucial ways. For example, it can never be spelled out on its own, and it always takes on the characteristics of whatever it occurs with. This is not the case for for example *u* in *iu*: it can be spelled out on its own, as second person, and it does not take on any characteristics of the *i* it occurs with in the inclusive. Ackema & Neeleman (To appear) argue that *a* is not a

(20)
a. *i*
b. *u*
c. *o*
d. *iu*
e. *io*
f. *uo*

The single atoms of person can be referred to by first, second and third person. The combination *iu* is referred to by the inclusive. If *io* and *uo* could be lexicalised in a similar way to inclusive, we would expect them to show the same paradigmatic distinctions the other persons do.

Consider the three-way singular-plural and minimal-augmented paradigms shown below. As explained in the introduction, singular-plural languages make a number division based on cardinality (i.e. one atom for singular, two for dual and more than two for plural) (21). Minimal-augmented languages on the other hand, make a number distinction based on presence or absence of associates (i.e. no associates for minimal, + one associate for unit-augmented and + more than one associates for augmented) (22). If *io* and *uo* were lexicalisable person distinctions, we would expect a 15-cell paradigm for singular-plural languages: 3 single atoms × 3 numbers + 3 combinations × 2 numbers (i.e. 3 persons for singular, 6 for dual and 6 for plural). For minimal-augmented languages, we would expect an 18-cell paradigm: 6 persons × 3 numbers (i.e. 6 for minimal, 6 for unit-augmented and 6 for augmented).

(21) Singular-plural

| SG | DU | PL |
|---|---|---|
| | *io* | *ioaaa...* |
| | *uo* | *uoaaa...* |
| | *iu* | *iuaaa...* |
| *i* | *ia* | *iaaa...* |
| *u* | *ua* | *uaaa...* |
| *o* | *oa* | *oaaa...* |

(22) Minimal-augmented

| MIN | UNIT | AUGM |
|---|---|---|
| *io* | *ioa* | *ioaaa...* |
| *uo* | *uoa* | *uoaaa...* |
| *iu* | *iua* | *iuaaa...* |
| *i* | *ia* | *iaaa...* |
| *u* | *ua* | *uaaa...* |
| *o* | *oa* | *oaaa...* |

These paradigms are not attested. The maximally distinct paradigm that is attested for singular-dual-plural, has 11 cells: 3 single atoms × 3 numbers + 1 combination × 2 numbers (i.e. 3 for singular, 4 for dual and 4 for plural). An example of this are Tümpisa Shoshone pronouns in (23). In distributing the personal pronouns over the paradigm, it is clear that there are no pronouns that could express the combinations *io* and *uo*. None of the existing forms can refer to a combination of hearer and other (*uo*) or of speaker and other (*io*). (Remember that e.g. *mümmü*, the second person plural, cannot refer to *uo* but refers to *ua* instead, as argued for in the previous section.)

number atom either, for which they make a convincing argument. However, they draw the conclusion that it must then be a part of the person system. Based on a.o. the distinctions between *a* and *i, u, o* just mentioned, it is unclear why they would assume this. Where exactly the referent *a* does enter the person and number derivation, is left for future research.

(23) Tümpisa Shoshone (Dayley 1979)

| SG | | DU | | PL | |
|---|---|---|---|---|---|
| | | io | * | ioaaa... | * |
| | | uo | * | uoaaa... | * |
| | | iu | *ta-ngku* | iuaaa... | *ta-mmü* |
| i | *nü* | ia | *nu-ngku* | iaaa... | *nü-mmü* |
| u | *ü* | ua | *mu-nku* | uaaa... | *mü-mmü* |
| o | (demonstr) | oa | (demonstr) | oaaa... | (demonstr) |

The maximally distinct paradigm attested for minimal-unit-augmented has 12 cells: 4 persons $\times$ 3 numbers (i.e. 4 for minimal, 4 for unit-augmented and 4 for augmented). These distribute over the paradigm as shown in (24) for Rembarrnga. There are pronouns for the four person distinctions *i, u, o, iu* in all three numbers but not for *io* and *uo* in any number.

(24) Rembarrnga (Cysouw 2003)

| MINIMAL | | UNIT-AUGMENTED | | AUGMENTED | |
|---|---|---|---|---|---|
| io | * | ioa | * | ioaaa... | * |
| uo | * | uoa | * | uoaaa... | * |
| iu | *yu̶kku̶* | iua | *ngakorru̶* | iuaaa... | *ngakorrbbarrah* |
| i | *ngu̶nu̶* | a | *yarru̶* | iaaa... | *yarrbbarrah* |
| u | *ku̶* | ua | *nakorru̶* | uaaa... | *nakorrbarrah* |
| o | *nawu̶/ ngadu̶* | oa | *barru̶* | oaaa... | *barrbarrah* |

To conclude: the sample does not have a single instance of a lexicalised *io* and *uo*. In the next sections (Sections 4-5), I show that this gap in lexicalisation is part of a broader generalisation on limitations on concept formation. Before that, I shortly explain the difference between a lexicalised combination and a syncretism for two separate person referents, as at first sight, syncretisms between *i* or *u* and *o* may look like lexicalisations of the combinations *io* and *uo*. I show that they are not.

### *3.3. Syncretism between* u *and* o *or* i *and* o

Some languages have a single pronoun to refer to both second and third person, or to both first and third person. I show that these are not instantiations of a lexicalised *uo* or *io* combination, since those would refer to the mereological sum of both atoms, whereas a syncretism refers to a disjunction.[13] An example is the pronoun for second and third person singular in Sanapaná, *hlejap*:

[13]If the pronouns are considered as accidentally homophonous instead of syncretic, there are two separate lexical entries that each refer to just one corner of the kite. As such, they do not form a problem for the kite to begin with. However, I follow analyses like Distributed Morphology (Halle & Marantz 1993) and Nanosyntax (Caha 2009) in that these are in facts syncretisms and therefore correspond to a single lexical entry that can be inserted in two syntactic structures. In this case, one form corresponds to the disjunction of two referents, as argued for in this section.

(25) Sanapaná (Harbour 2015)

| | SG | PL |
|---|---|---|
| iu | | *enenko'o* |
| i | *ko'o* | *enenko'o* |
| u | *hlejap* | *hlengap* |
| o | *hlejap* | *hlengap* |

The interpretation of inclusive shows that lexicalised person combinations are mereological sums (I return to this distinction in Section 5.1). This means that the inclusive refers to the sum of speaker and hearer, *i* and *u*, but not to their disjunction, *i* or *u*. Consider again Tümpisa Shoshone, which has an established lexicalised combination: *tammü* for *iu*. Crucially, this pronoun can only refer to the mereological sum of speaker and hearer, not to either the speaker or the hearer, as the example below shows. If reference to either the speaker or the hearer is needed, a first or second person pronoun is used.

(26) a. **Tammü** piiya hipi"- tükin- tu'ih
**we.incl** beer drink- start- will
'**We**'ll start to drink beer.' (Dayley 1979:14)

b. **Nü** tattanguhi punikka
**I** men.DU.O see
'**I** see two men.' (Dayley 1979:224)

c. **Mümmü** ha namo'okoitsoimmaa?
**you-PL** Q handwashed
'Did **you** wash your hands?' (Dayley 1979:325)

In contrast, syncretic pronouns are not mereological sums but disjunctions instead. They are merely pronouns that can be used for more than one concept, as in non-inclusive languages like English. *We* can be used to refer to the inclusive person or to the first person plural, depending on the context. Consider the examples below, where *we* in (27a) refers to the speaker and the hearer (and possibly associates) *iu(a)*, but *we* in (27b) does not refer to the hearer, therefore corresponding to *ia* (i.e. the speaker and his associates). This shows that in syncretism, there is no obligatory mereological sum, as opposed to lexicalised combinations.

(27) Context: You are talking to someone from a different university whom you just met at a conference:

a. Don't you think **we** get excellent coffee at this conference?

b. In my office in Brussels, **we** have excellent coffee.

Now consider languages like Sanapaná in (25), where one pronoun, *hlejap* (and *hlengap* in plural), can refer to *u* or to *o*. *Hlejap* and *hlengap* could mistakenly be analysed as lexicalised *uo* combinations; however, I show that they are not. They behave like the English *we* rather than the Tümpisa Shoshone *tammü*, and are therefore not lexicalised combinations but rather syncretisms. The Sanapaná sentences below show that the pronoun *hlejap* refers to either second or third person, depending on the context, but never to the sum of both, as a plural translation 'you and him' for *hlejap* is simply not found in the data. As such, this is not the mereological sum *uo* and therefore not a lexicalised combination.

(28) a. **Hlejap** metko patakon ap- angok.
**2sg/3sg** NEG money 2sg/3sg- POSS
'**He** doesn't have money.'
b. Ta'asak akjehlna ap- ta= o **hlejap**?
which fruit 2sg/3sg- eat- Q **2sg/3sg**
'Which fruit did **you** eat?' (Harbour 2016:62)

The same holds for languages with a first and third person syncretism. These are not lexicalised combinations, but rather syncretisms of two different referents.
Consider for example the Spanish imperfect person suffixes:[14]

(29) a. habla- ba- **ø**
speak- PAST- **1/3SG**
'**I/she/he** spoke'
b. habla- ba- **s**
speak- PAST- **2SG**
'**You** spoke' (Cysouw 2003:46)

Since Spanish does not require pronouns to be present, the sentence in (29a) is ambiguous: it could refer either to a first person or to a third person, but necessarily not to a combination of both, as I am not aware of such a (necessarily plural) interpretation. As such, this is not an instance of a lexicalised *io* combination.

The data from the sample demonstrates that from all logically possible person atom combinations, *io* and *uo* cannot be lexicalised in natural language. In what follows, I show how this limitation on the lexicalisation of person is part of a broader generalisation on concept formation in closed lexical fields, as observed in the kite framework. First, I discuss the framework itself (Section 4) to then explain how it applies to person and predicts exactly the pattern I have just shown (Section 5).

## *4. The kite framework*

In this section, I briefly explain the basis of the kite framework, and exemplify it with the lexical field of quantifiers. Then, I discuss the prediction this framework makes about lexicalisation in closed lexical fields: the CONCEPT FORMATION CONSTRAINT (Jaspers 2012; Seuren & Jaspers 2014; Roelandt 2016).

### *4.1. The logical square and hexagon*

The kite framework represents the concepts of closed lexical fields and the logical relations that hold between them in geometrical figures (Seuren & Jaspers 2014). This is a tradition that goes back to Aristotle's square of opposition, shown in Fig. 2. The square can be split up into two

[14] I am not aware of pronouns that show this syncretism in the singular as it is an uncommon pattern. Cysouw (2003) does list instances of plural pronouns with this syncretism, e.g. Bagirmi. The singular examples illustrate the point made more clearly, even though the plural syncretism shows the same kind of disjunction as the singular does. Therefore, this example is one of agreement markers rather than independent pronouns.

subuniverses: a positive one, with a universal affirmative (A) and a particular affirmative (I) (for the 'A' and 'I' in *affirmo*); and a negative one, with a universal negative (E) and a particular negative (O) (for the 'E' and 'O' in *nego*). The square can model closed lexical fields such as the quantifiers in Fig. 3.

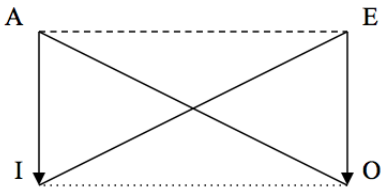


*Figure 2*: *Aristotle's square of opposition*

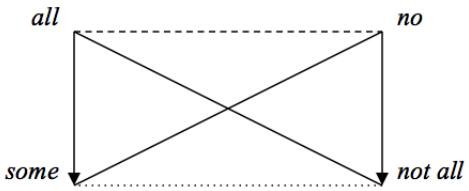


*Figure 3*: *The square of opposition: quantifiers*

Between the corners, certain logical relations hold (Jaspers 2012):

(30) a. Entailment (arrows):
  (i) If A is true, then I is also true.
  (ii) If E is true, then O is also true.
  (iii) E.g.: If *all* Belgians love chocolate, then *some* Belgians love chocolate.

b. Contradiction (full lines):
  (i) A and O cannot both be true or both be false at the same time.
  (ii) I and E cannot both be true or both be false at the same time.
  (iii) E.g. '*All* Belgians love chocolate' and '*No* Belgians love chocolate' cannot be true or false at the same time.

c. Contrariety (dashed line):
  (i) A and E cannot both be true but they can both be false at the same time.
  (ii) E.g. '*All* Belgians love chocolate' and '*No* Belgians love chocolate' cannot be true at the same time, but they can be false at the same time, i.e. when only *some* Belgians love chocolate.

d. Subcontrariety (dotted line):
  (i) I and O can both be true but not both be false at the same time.
  (ii) E.g. '*Some* Belgians love chocolate' and '*Not all* Belgians love chocolate' can both be true at the same time, but they cannot be false at the same time.

It has long been noted that the quantifier *some* can be ambiguous. Consider the following sentences:

(31) If some students pass the test, I'll treat them with chocolates.

(32) Some students have passed the test.

*Some* in (31) can also refer to *all*, since I will still treat the students to chocolate if all of them pass the test. It is therefore also called the 'inclusive *some*', since it can include *all* (Roelandt 2016:108). If all students passed the test, the sentence in (32) would be false, since here *some* does not mean *all*. This is called the 'exclusive *some*'.

Jacoby (1950, 1960), Sesmat (1951) and Blanché (1952) attribute this difference to the existence of two different quantifiers *some*: an inclusive and an exclusive one (Jaspers 2012; Seuren & Jaspers 2014). Therefore, they extend the square to a hexagon (Figs. 4 and 5), which now

includes both kinds of *some* (in I and Y) and contradictories to each of them (in E and U respectively).[15]

Nine more relations run to and from these new corners: a contradiction between Y and U, four extra entailment relations, a triangle of contrariety (A-Y-E) and a triangle of subcontrariety (I-U-O). The triangle of contrariety is made up of concepts that consist of the conjunction of their adjacent corners (e.g. *some* in the Y-corner: '*some* and *not all*') and the triangle of subcontrariety is made up of corners that consist of the disjunction of their adjacent corners (e.g. the I-corner: '*all* or *some*$_{excl}$').

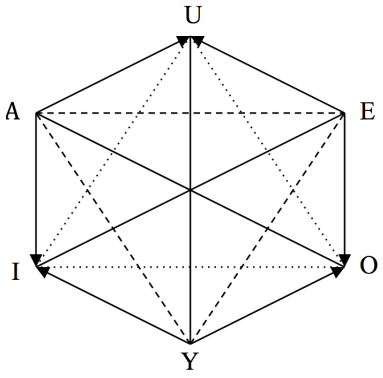


*Figure 4*: *Hexagon*

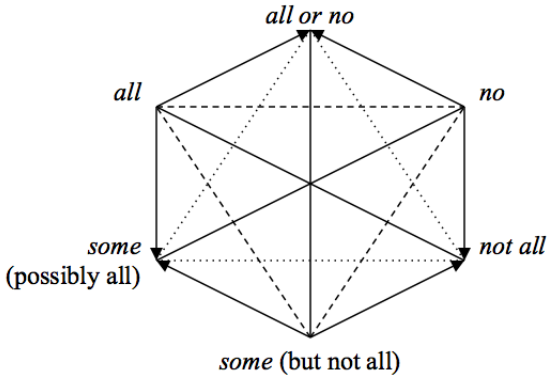


*Figure 5*: *Quantifier hexagon*

This hexagon now shows all the distinctions in the conceptual field of quantifiers and the logical relations that hold between the concepts.

### *4.2. The Concept Formation Constraint*

Jaspers (2005, 2012) and Seuren & Jaspers (2014) note that not all of the corners on the hexagons can be lexicalised: both the O- and the U-corner do not receive a simplex lexicalisation in natural language.[16] This is generalised in the CONCEPT FORMATION CONTRAINT.

(33) THE CONCEPT FORMATION CONSTRAINT (CFC)
When mapping a closed lexical field onto the logical hexagon, the O- and U-corners never receive a simplex lexicalisation.

For quantifiers, this means that we cannot express *not all* (O) and *all or no* (U) as simplex lexical items. Omitting these O- and U-corners from the hexagons in Figs. 4 and 5 results in the kite.

[15]Grice (1975) attributes the two possible readings of *some* to pragmatic implicature, rather than to a semantic difference. In contrast, the approach in terms of the hexagon developed by Jacoby, Sesmat and Blanché takes the distinction between inclusive and exclusive *some* to be a semantic one (see Seuren & Jaspers 2014:620 for their arguments). Since this issue is orthogonal to my concerns, I do not discuss it any further here.

[16]For the O-corner, this was already observed by Thomas Aquinas and later generalised by Blanché (1953, 1966).

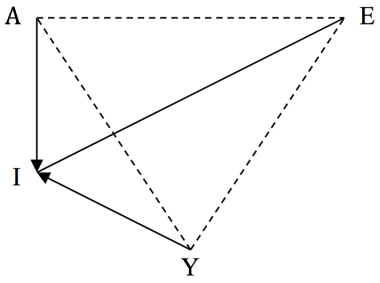


*Figure 6*: *Kite*

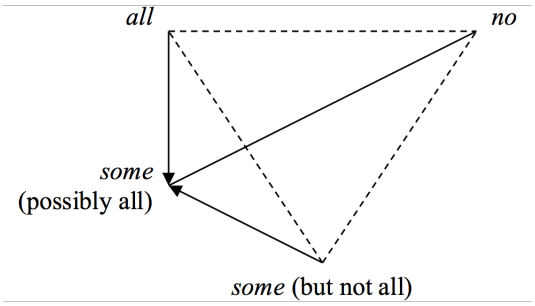


*Figure 7*: *Quantifier Kite*

The kite accounts for all the lexicalisable distinctions in the sense that it shows all the distinctions attested within the conceptual field of quantifiers and correctly rules out the unlexicalisable concepts. There is no simplex lexical item *allno* to express the disjunction *all or no*, neither is there a *nall* in natural language to express *not all*.

The CFC has been shown to apply to quantifiers as demonstrated above, to the predicate calculus logical operators *and, or, nor* and to binary lexical fields with an overarching hyperonym that can be split up into subdomains, for example 'human' as an overarching lexical predicate for 'man' and 'woman' by Seuren & Jaspers (2014). It has also been applied to colour perception terms in Jaspers 2012 and extended for adjectives by Roelandt (2016). In all these cases, when mapping the lexical field onto the hexagon, the same two corners do not receive a simplex lexicalisation. Details can be found in the works cited.

In what follows, I argue that this same CONCEPT FORMATION CONSTRAINT holds for person (as already hinted at in Seuren & Jaspers 2014) and show that the absence of a person morpheme for *io* and *uo* is part of this broader generalisation on lexicalisation limitations.

## *5. The person kite*

In this section, I explain how the logical relations mentioned above apply to person in Section 5.1. Section 5.2 presents how combining these relations with the person atoms results in the person kite and how the person concepts end up in their respective corners. Applying the CFC to this derived kite, will correctly predict the inclusive *iu* to be the only lexicalisable person atom combination since *io* and *uo* will be ruled out as lexicalisable in natural language.

### *5.1. The mereological relations of the person kite*

As explained in Section 4.1, the hexagon has relations of entailment, contradiction, contrariety and subcontrariety between its vertices. These are logical relations that hold between propositions. For quantifiers, for example, the entailment relation *all* $\subseteq$ *some* really holds between propositions of the sort *All Belgians like chocolate* and *Some Belgians like chocolate*. Person atoms and their combinations, on the other hand, do not correspond to propositions but to person referents. Therefore, it is unclear how the logical relations described above could apply. For example, for an entailment to hold between *i* and *iu*, these person atoms would need to be propositions: e.g. 'x is a speaker' for *i*. However, for *iu* this would correspond with 'x is

a participant'. Under such an interpretation, there is an entailment from *i* to *iu*, since being a speaker entails being a participant. However, in this case *iu* would correspond to *i or u* (since being a participant does not mean being both a speaker and a hearer, but being either a speaker or a hearer) and that does not correspond to the reference of the inclusive person.[17] Instead, the inclusive refers necessarily to the mereological sum of the speaker and the hearer. Therefore, for person, I take the relations to be of a mereological nature.

Mereology is the theory of parthood relations (Varzi 2016). This means that rather than dealing with entailment relations between logical concepts and propositional truth values, a mereology deals with parthood relations between atoms in a universe and their mereological sums (such as cake being the mereological sum of flour, butter, eggs and sugar). For example, for person, both *i* and *u* are parts of *iu* and conversely, *iu* is the mereological sum of *i* and *u* ($i \oplus u$).

The relations in logical and mereological hexagons show a very clear isomorphism. The corresponding relations (and the labels used for both in this paper) are shown in the table in (34). (35) shows the logical relations already exemplified in Section 4.1 and (36) explains the corresponding mereological relations.

(34) The corresponding logical and mereological relations

| Logical | Mereological | Label |
|---|---|---|
| Entailment | Parthood | $\subseteq$ |
| Contradiction | Exhaustive complementarity | CD |
| Contrariety | Non-exhaustive complementarity | C |
| Subcontrariety | Exhaustive overlap | SC |

(35) Logical relations:

a. **Entailment**: A entails B iff when A is true, B is also true.
b. **Contradiction**: A and B are contradictory iff they cannot both be true or both be false at the same time.
c. **Contrariety**: A and B have a contrariety relation iff they cannot both be true but they can both be false at the same time.
d. **Subcontrariety**: A and B have a subcontrariety relation iff they can both be true but not both be false at the same time.

(36) Mereological relations:

a. **Parthood**: A $\subseteq$ B iff A is a part of B.
b. **Exhaustive complementarity**: A and B are exhaustive complements iff their intersection is empty (i.e. they have no overlapping elements) and their union equals the universe.
c. **Non-exhaustive complementarity**: A and B are non-exhaustive complements iff their intersection is empty but their union does not equal universe.
d. **Exhaustive overlap**: A and B show exhaustive overlap iff their intersection is non-empty (i.e. they have at least one overlapping atom); their union equals the universe, and neither is part of the other.

[17] Another possible entailment is the following: 'we (*iu*) like coffee' entails 'I (*i*) like coffee'. This, however, is an entailment that does not hold with collective predicates, i.e. 'we gathered in the hall' does not entail 'I gathered in the hall'. (I would like to thank Guido Vanden Wyngaerd for pointing this out to me.) In short, entailments, however construed, are problematic for the inclusive, and need to be replaced by mereological relations.

The isomorphism between these relations becomes clear when employing bitstrings to represent the corners of the hexagon (Smessaert 2009; Jaspers 2012; Roelandt 2016). In general, any hexagon can be described as composed of three atoms (*i, u* and *o* for person and *all, some* and *no* for quantifiers) and their combinations. The number of possibilities equals $2^3 = 8$. Discounting 000 and 111, this yields the six corners of the hexagon. Therefore, the bitstrings are of length three with every position in the bitstring corresponding to exactly one atom: a 1-value indicates the presence of this atom and a 0-value the absence. Combining the three atoms leaves us with eight logical possibilities, of which I will not discuss the absence of all atoms, 000, nor the presence of all atoms, 111 (as motivated in Footnote 2).[18]

(37) Logical possibilities with bitstrings

| Bits | Quantifiers | Person |
|---|---|---|
| 100 | *all* | *i* |
| 010 | *some* (*some*$_{excl}$) | *u* |
| 001 | *no* | *o* |
| 110 | *all or some* (*some*$_{incl}$) | *iu* |
| 101 | *all or no* | *io* |
| 011 | *some or no* (*not all*) | *uo* |

These bitstrings can be used to formalise the logical and mereological relations explained above, and show how they are parallel (Smessaert 2009). For each mereological relation I mention the equivalent logical relation and the abstract formula for both, with $\phi$ and $\psi$ representing the concepts on the vertices of the hexagon. In these formulas, intersection ($\cap$) selects only the atoms (i.e. 1-values) both items have in common, whereas union ($\cup$) selects all the atoms (i.e. 1-values) present in either or both of the items. For example, 100 $\cap$ 010 = 000 and 100 $\cup$ 010 = 110. Underneath the general formulas, I show how this applies to quantifiers in a. and to person in b. I also show how both are generalised in the same way by means of the bitstrings in c.

(38) Parthood and entailment ($\subseteq$):
$\phi \subseteq \psi$ iff $\phi \cap \psi = \phi$ and $\phi \cup \psi = \psi$
a. *all* $\subseteq$ *some*$_{incl}$ because *all* $\cap$ *some*$_{incl}$ = *all* and *all* $\cup$ *some*$_{incl}$ = *some*$_{incl}$
b. *i* $\subseteq$ *iu* because *i* $\cap$ *iu* = *i* and *i* $\cup$ *iu* = *iu*
c. 100 $\subseteq$ 110 because 100 $\cap$ 110 = 100 and 100 $\cup$ 110 = 110

(39) Exhaustive complementarity and contradiction (CD):
$\phi$ CD $\psi$ iff $\phi \cap \psi$ = 000 and $\phi \cup \psi$ = 111
a. *some*$_{incl}$ CD *no* because *some*$_{incl}$ $\cap$ *no* = 000 and *some*$_{incl}$ $\cup$ *no* = 111
b. *iu* CD *o* because *iu* $\cap$ *o* = $\varnothing$ and *iu* $\cup$ *o* = *iuo*
c. 110 CD 001 because 110 $\cap$ 001 = 000 and 110 $\cup$ 001 = 111

(40) Non-exhaustive complementarity and contrariety (C):
$\phi$ C $\psi$ iff $\phi \cap \psi$ = 000 and $\phi \cup \psi \neq$ 111
a. *all* C *no* because *all* $\cap$ *no* = 000 and *all* $\cup$ *no* $\neq$ 111

[18]Note that whether or not 111 and 000 receive simplex lexicalisation can differ from lexical field to lexical field. For colour for example, 000 corresponds to black and 111 to white. These concepts are included in the hexagon with a third dimension: both are points in the centre, one sticking out on top, the other at the bottom Jaspers (2012). This is likely also the case for person.

b. *i* C *o* because $i \cap o = \varnothing$ and $i \cup o \neq iuo$
c. 100 C 001 because $100 \cap 001 = 000$ and $100 \cup 001 \neq 111$

(41) Exhaustive overlap and subcontrariety (SC):
$\phi$ SC $\psi$ iff $\phi \cap \psi \neq 000$; $\phi \cup \psi = 111$ ; $\phi \not\subseteq \psi$; $\phi \not\supseteq \psi$

a. $some_{incl}$ SC *not all* because $some_{incl} \cap$ *not all* $\neq 000$; $some_{incl} \cup$ *not all* = 111; $some_{incl} \not\subseteq$ *not all*; $some_{incl} \not\supseteq$ *not all*
b. *iu* SC *uo* because $iu \cap uo \neq \varnothing$; $iu \cup uo = iuo$; $iu \not\subseteq uo$; $iu \not\supseteq uo$
c. 110 SC 011 because $110 \cap 011 \neq 000$; $110 \cup 011 = 111$; $110 \not\subseteq 011$; $110 \not\supseteq 011$

To illustrate, consider the hexagons below with the bitstrings added to the corners.[19] (I explain exactly how the person hexagon is derived in the next section.)

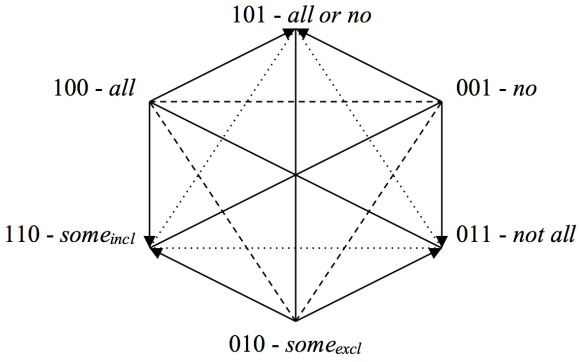


*Figure 8*: *Quantifier hexagon with bitstrings*

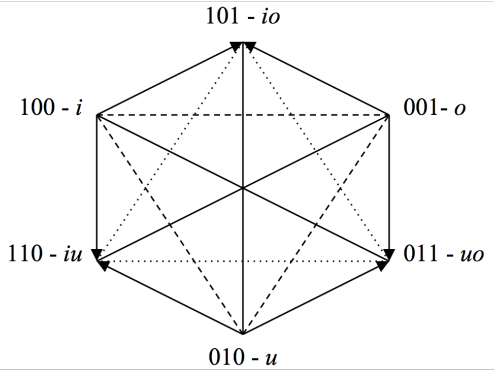


*Figure 9*: *Person hexagon with bitstrings*

This shows that even though logical and mereological relations apply to different kinds of concepts (i.e. propositions vs. referents), they show an isomorphism, which can be brought out with the bitstring formalism.

The crucial difference between logical and mereological lexical fields relates to the I-O-U corners (see Fig. 4 for the hexagon with letters representing the corners). In logical systems, these make up the triangle of subcontrariety and correspond to the disjunction of their adjacent corners. For example, for quantifiers, the I-corner $some_{incl}$ corresponds to $some_{excl}$ *or all* (Fig. 10). For mereologies, these corners (making up the triangle of exhaustive overlap) are the mereological sums ($\oplus$) of their adjacent corners (Fig. 11). This means that for the person hexagon I derive below, *iu* corresponds to $i \oplus u$.

[19] Note that which atom receives which 1-bit in the bitstring is arbitrary and does not affect the relations explained below. However, which concept ends up in which concept of the hexagon is not arbitrary, and will be explained in Section 5.

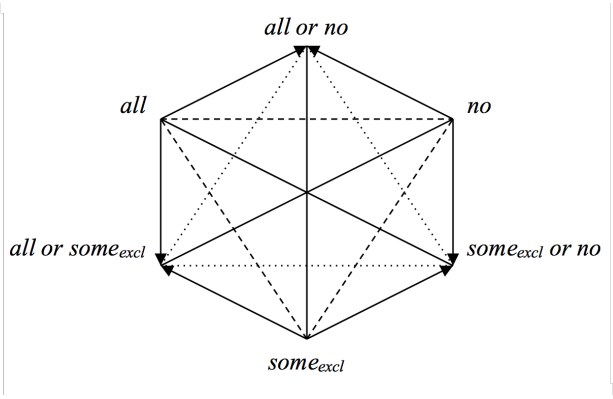


*Figure 10*: *Quantifier hexagon: disjunction*

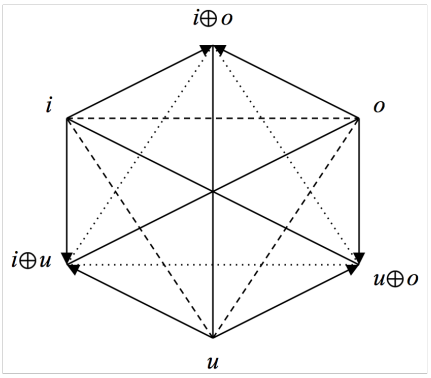


*Figure 11*: *Person hexagon: mereological sum*

Despite this difference, mereological hexagons are also subject to the CFC, just like logical ones.[20]

### *5.2. Deriving the person kite*

Now that we have seen how the CFC applies to closed lexical fields and that person is a mereological lexical field, I show exactly how the person kite is derived and why the concepts end up in their respective corners.

Seuren & Jaspers (2014) show how from one single entailment relation between two concepts (e.g. *all* $\subseteq$ *some*$_{incl}$), a hexagon, and by virtue of the CFC also a kite, automatically follows. (To see how the parallel with any other lexical field works in detail, see Seuren & Jaspers 2014:612-613;626-627.) For a mereology, the hexagon is derived from a parthood relation: $i \subseteq iu$.

Here I explain the step by step derivation of the person kite and illustrate with Fig. 12.

- **Step 1**: The derivation of the person hexagon starts with the parthood relation between *i* and *iu*: *i* is a part of *iu* because it is contained in *iu*.

- **Step 2**: From any parthood relation, two exhaustive complementarity relations follow. One is between *iu* and *o*: *iu* and *o* are exhaustive complements because they have no overlapping elements and their union equals the universe (i.e. *iuo*). The other is between *i* and *uo*: *i* and *uo* also have no overlapping elements but their union equals the universe.

- **Step 3**: These two new corners stand in a parthood relationship: *o* is a part of *uo*. There is also a non-exhaustive complementarity relationship between the single atoms, *o* and *i*: they are non-exhaustive complements because they have no overlapping elements but their union does not equal the universe (i.e. *u* is still a possibility in the universe). Finally, there is an exhaustive overlap relation between the combinations *iu* and *uo*: they overlap exhaustively because their union equals the universe *iuo*, their intersection is non-empty because it contains *u*, and they are no subsets because *i* and *o* are complementary. The result is the mereological equivalent of the square of opposition.

[20]This has been demonstrated for the mereology of colour, see Jaspers 2012 for a full discussion.

- **Step 4**: The atoms *i* and *o* each have their own corner, but the atom *u* in *iu* and *uo* does not. Adding this final single atom, *u*, completes the universe and thereby finishes the triangle of non-exhaustive complementarity. There are also two more parthood relations: from *u* to *iu* and from *u* to *uo*.
- **Step 5**: There is now a new complementarity relation between *u* and *io*. There are also two more parthood relations, from *i* and *o* to *io*, and the triangle of exhaustive overlap is finished.
- **Step 6**: At this point, taking out the unlexicalisable corners *io* and *uo* turns the hexagon into the kite.

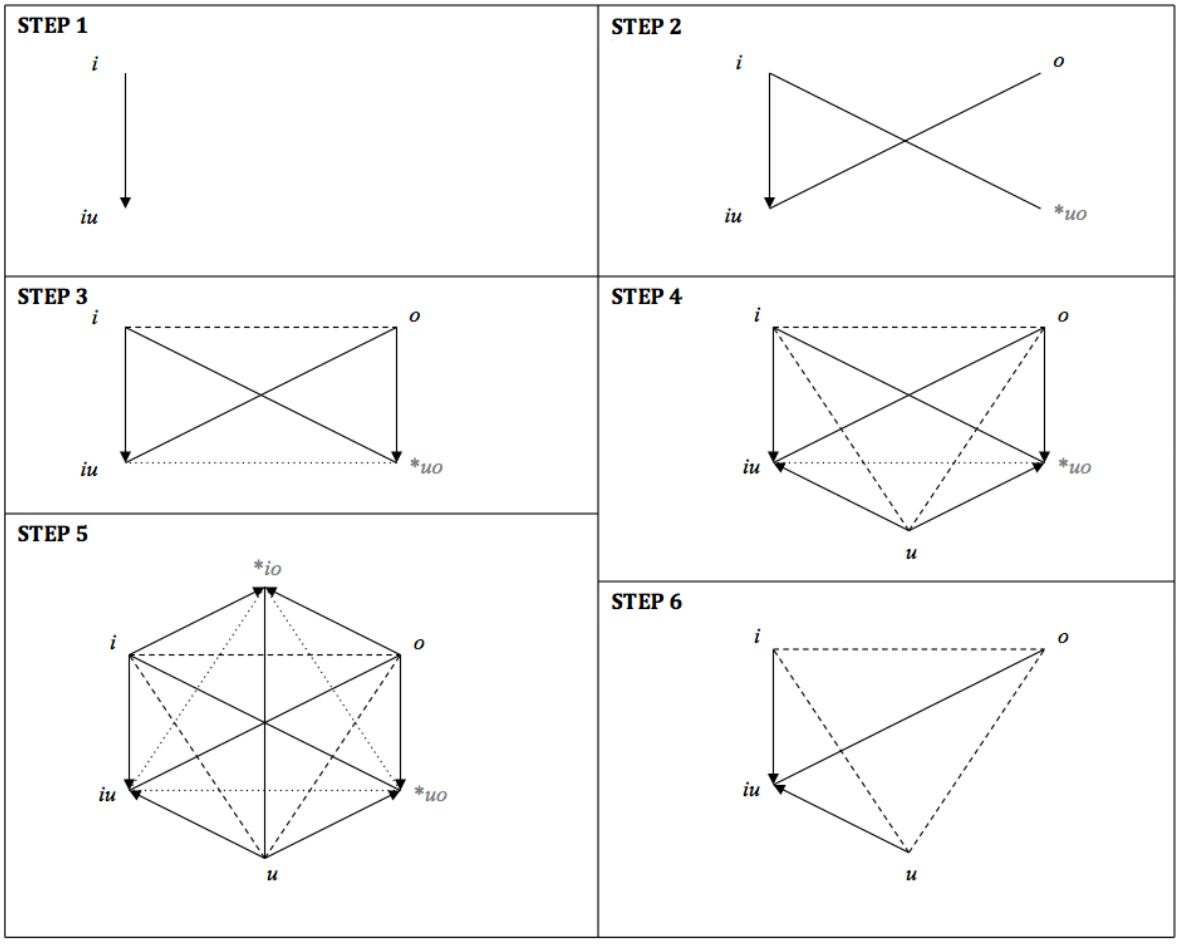


*Figure 12*: *The derivation of the person kite step by step*

Associates do not occupy a position in the hexagon, the reason for this being that they are not person atoms (as explained in Footnote 12).

Note that there are other parthood relations that can derive a hexagon, e.g. $o \subseteq io$. However, apart from representing the mereological relations, the hexagon also shows a divide between a positive universe (A-Y-I, remember: A and I in *affirmo* as explained in Section 4.1) and a negative one (U-E-O, E and O in *nego*) (Jaspers 2012:239; Seuren & Jaspers 2014:622). In the quantifiers for example, the only negative quantifier *no* resides both in the U and E corner, and in

the O-corner we find a contradiction of a positive quantifier. For colour, the colours in the positive universe are often considered as warm colours, whereas the ones in the negative universe are often considered as cold. Other parthood relations amongst person atoms, e.g. between *o* and *io*, may also provide a mereological hexagon, but this will not follow the positive-negative divide. Third person, *o*, belongs to the negative side of the person universe since, as opposed to the other person atoms, it does not refer to a speech act participant and, morphologically, it often differs from the other persons in for example selecting different number morphology or being syncretic with demonstratives (a.o. Silverstein 1976; Nevins 2007; Bobaljik 2008).

Therefore, the only two viable parthood relations to start from are *i* or *u* $\subseteq$ *iu*. Both result in a mereological hexagon with the participants on the positive side and non-participants on the negative side. So far, I have not come across any convincing arguments to start with one rather than the other. Note that this choice does not affect the predictions made for person lexicalisation: in both cases, *io* and *uo* will end up in the O- and U-corners and therefore be predicted to be unlexicalisable.

I illustrate this with the example from Section 3.2, Tümpisa Shoshone (23). The person morphemes available correspond to the corners of the kite and there are no morphemes for the concepts in the O- and U-corners (Fig. 13).

Note that the generalisation on the lexicalisation of person made in this paper concerns person morphemes and not full person and number markers. As such, the corners in the hexagons above represent only the person morphemes. For fusional paradigms like English (i.e. paradigms that do not have separate person and number morphemes but express both in just one morpheme) the kite looks more filled: the singular, plural, different genders, etc., are placed together in their relevant person corner. However, in fusional paradigms as well there is no pronoun to express a combination *io* or *uo* (Fig. 14).

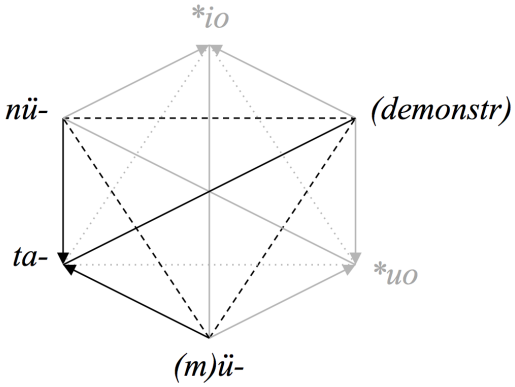


*Figure 13*: *Hexagon for Tümpisa Shoshone*

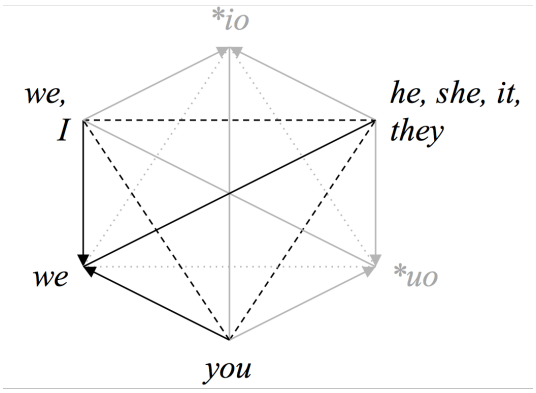


*Figure 14*: *Hexagon for English*

## *6. Conclusion*

I have shown that plural is not a case of '+ third person' since morphologically, no languages share a morpheme for third person and plural and semantically, the reference of a plural pronoun cannot include a third person (Ackema & Neeleman To appear). Instead, plural is a case of '+ associates'. In order for a third person to be included in the reference of a plural pronoun, it first needs to be established as an associate.

Also, syncretism between second and third person or first and third person is not an instantiation of a lexicalised person atom combination, since syncretic pronouns refer to only one of the persons it can express (depending on the situation), i.e. to $u \cup o$ and $i \cup o$, and not to the mereological sum of these, $u \oplus o$ and $i \oplus o$.

From all logically possible person referents and their combinations, two groups can never be lexicalised: *io* and *uo*. Exactly the gap in the lexicalisation of person that *io* and *uo* leave is predicted by the kite framework. Its CONCEPT FORMATION CONSTRAINT makes a prediction about lexicalisation in closed lexical fields: when representing the concepts of these lexical fields in a logical hexagon, the same two corners are never lexicalised in natural language. Applying this framework to person shows that the unlexicalisability of *io* and *uo* is part of this broader generalisation on limitations of concept formation.

Jolijn Sonnaert
KU Leuven - CRISSP
www.crissp.be
jolijn.sonnaert@kuleuven.be

*Acknowledgements*

Most of all, I would like to thank Guido Vanden Wyngaerd for his extensive advice and helpful comments for the research conducted for and the writing of this paper. My thanks also go out to Dany Jaspers for the enthusiastic discussions that have helped me figure out the details of the framework. I am also grateful to Karen De Clercq, Jeroen van Craenenbroeck, Pieterjan Benoit, Cora Pots, Leen Sevens and an anonymous reviewer for their feedback during the writing process.

*References*

Ackema, P. & A. Neeleman (To appear). *Features of Person*.
Baerman, M., D. Brown & G. G. Corbett (2005). *The Syntax-Morphology Interface*. Cambridge University Press, Cambridge.
Blanché, R. (1952). Quantity, modality and other kindred systems of categories. *Mind* 61:243, pp. 369–375.
Blanché, R. (1953). Sur l'opposition des concepts. *Theoria* :19, pp. 89–130.
Blanché, R. (1966). *Structures intellectuelles: Essai sur l'organisation systématique des concepts*. J. Vrin, Paris.
Bobaljik, J. D. (2008). Missing persons. *The Linguistic Review* 25, pp. 203–230.
Caha, P. (2009). *The Nanosyntax of Case*. Ph.D. thesis, University of Tromsø.
Cysouw, M. (2003). *The Paradigmatic Structure of Person Marking*. Oxford University Press, New York, 2009 edn.
Daniel, M. (2005). Understanding inclusives. Filimonova, E. (ed.), *Clusivity: Typology and case studies of the the inclusive–exclusive distinction*, John Benjamins Publishing Co., pp. 3–48.
Dayley, J. P. (1979). *Tümpisa (Panamint) Shoshone Grammar*. University of California Press, Berkely.
Dryer, M. S. & M. Haspelmath (eds.) (2013). *The World Atlas of Language Structures Online*. Max Planck Institute for Evolutionary Anthropology, Leipzig.
Forchheimer, P. (1953). *The category of person in language*. Walter de Gruyter & Co, Berlin.
Ford, K. & D. Ober (2004). A sketch of Kalaw Kakaw Ya. Romaine, S. (ed.), *Language in Australia*, Cambridge University Press, Cambridge, chap. 7, pp. 118–142.
Grice, P. H. (1975). Logic and conversation. Cole, P. & J. L. Morgan (eds.), *Syntax and semantics*, Academic Press, New York, vol. 3, pp. 41–58.

Halle, M. & A. Marantz (1993). Distributed Morphology and the pieces of inflection. Hale, K. & S. J. Keyser (eds.), *The View from Building*, MIT Press, vol. 20, pp. 111–176.

Harbour, D. (2015). Poor pronoun systems and what they teach us. Krämer, M., S.-I. Ronai & P. Svenonius (eds.), *Nordlyd*, University of Tromsø, Tromsø, vol. 40.

Harbour, D. (2016). *Impossible Persons*. MIT Press, Cambridge, Massachusetts.

Harley, H. & E. Ritter (2002). Person and number in pronouns: A feature-geometric analysis. *Language* 78:3, pp. 482–526.

Jacoby, P. (1950). A triangle of opposites for types of propositions in aristotelian logic. *The New Scholasticism* :24, pp. 32–56.

Jacoby, P. (1960). Contrariety and the triangle of opposites in calid inferences. *The New Scholasticism* :34, pp. 141–169.

Jaspers, D. (2005). *Operators in the Lexicon*. Ph.D. thesis, Universiteit Leiden LOT, Utrecht.

Jaspers, D. (2012). Logic and colour. *Logica Universalis* 6, pp. 227–248.

Moskal, B. (2014). The role of morphological markedness in inclusive/exclusive pronouns. *Annual Meeting of the Berkeley Linguistics Society*, pp. 354–368, URL http://journals.linguisticsociety.org/proceedings/index.php/BLS/article/view/3148/2867.

Nevins, A. (2007). The representation of third person and its consequences for person-case effects. *Natural Language and Linguistic Theory* 25, pp. 273–313.

Ray, S. H. (1907). *Reports of the Cambridge anthropological expedition to Torres Straits*, vol. 3. Cambridge University Press, Cambridge.

Roelandt, K. (2016). Most *or the art of compositionality: Dutch* de/het meeste *at the syntax-semantics interface*. Ph.D. thesis, KU Leuven, LOT, Brussels.

Round, E. & L. Stirling (2015). Universals of split argument coding and morphological neutralization: Why Kala Lagaw Ya is not as bizarre as we thought. *Australian Journal of Linguistics* 35:3, pp. 251–281.

Sesmat, A. (1951). *Logique: Les Raisonnements. La Logistique.* Hermann.

Seuren, P. A. M. & D. Jaspers (2014). Logico-cognitive structure in the lexicon. *Language* 90:3, pp. 607–643.

Silverstein, M. (1976). *Hierarchy of features and ergativity*, Australian Institute of Aboriginal Studies, pp. 112–171.

Smessaert, H. (2009). On the 3D visualisation of logical relations. *Logica Universalis* 3, pp. 303–332.

Varzi, A. (2016). Mereology. *The Stanford Encyclopedia of Philosophy* URL http://plato.stanford.edu/archives/spr2016/entries/mereology/.